

METHOD AND SYSTEM OF A NETWORK LINK ADAPTER

BACKGROUND

[0001] When a network adapter interfaces different network links, the speed of communication on the network may be undesirably limited by the slowest link. Therefore, methods and systems that increase the speed of communication between different network links are desirable.

SUMMARY

[0002] In at least some embodiments, a system may comprise a network link adapter, wherein the network link adapter is configurable to detect when a first device transmits data targeting a second device and write the data targeting the second device to an address of a memory. The system further comprises a monitoring unit coupled to the network link adapter, the monitoring unit detects when the data is written to the address of the memory and forwards the data to the second device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] For a detailed description of exemplary embodiments of the invention, reference will now be made to the accompanying drawings in which:

[0004] Figure 1 illustrates a block diagram of a system in accordance with embodiments of the invention;

[0005] Figure 2 illustrates a flow of data in the system of Figure 1 in accordance with embodiments of the invention; and

[0006] Figure 3 illustrates a method in accordance with embodiments of the invention.

NOTATION AND NOMENCLATURE

[0007] Certain terms are used throughout the following description and claims to refer to particular system components. As one skilled in the art will appreciate, computer companies may refer to a component by different names. This document does not intend to distinguish between components that differ in name but not function.

[0008] In the following discussion and in the claims, the terms “including” and “comprising” are used in an open-ended fashion, and thus should be interpreted to mean “including, but not limited to... .” Also, the term “couple” or “couples” is intended to mean either an indirect or direct electrical connection. Thus, if a first device couples to a second device, that connection may be through a direct electrical connection, or through an indirect electrical connection via other devices and connections.

DETAILED DESCRIPTION

[0009] The following discussion is directed to various embodiments of the invention. Although one or more of these embodiments may be preferred, the embodiments disclosed should not be interpreted, or otherwise used, as limiting the scope of the disclosure. In addition, one skilled in the art will understand that the following description has broad application, and the discussion of any embodiment is meant only to be exemplary of that embodiment, and not intended to intimate that the scope of the disclosure is limited to that embodiment.

[0010] As will be disclosed below, various embodiments of the invention improve the speed of communication in a multi-link network. In at least some embodiments, the speed of network communication may be improved by detecting when a first device transmits data targeted for a second device and providing an alternate route (separate from a network link that couples the first and second devices) for the targeted data to be received by the second device. The second device may thereby simultaneously receive data from the first device via the alternate route and output data to the first device via the network link.

[0011] Figure 1 illustrates a system 100 in accordance with embodiments of the invention. As shown in Figure 1, the system 100 may comprise a first device 102 and a second device 104 coupled to a network link adapter 106 via a first

interface 120 and a second interface 122, respectively. In at least some embodiments, the first device 102 may comprise a switching node coupled via a network connection to one or more remote computer systems and/or devices (not shown). Therefore, the first device 102 may function as a switching device that permits data to be transmitted between the second device 104 and any remote computer systems coupled to the first device 102. The second device 104 may comprise a programmable device such as a field programmable gate array (FPGA). In some embodiments, the second device 104 may be programmed to process and/or combine predetermined data types (e.g., image data) received from remote computers systems coupled to the second device 104 via the first device 102 and the network node adapter 106 as described above. In addition, the second device 104 may process and/or combine data received via the first device 102 with data received from a local computer system coupled to the second device 104.

[0012] The first device 102 and the second device 104 may be embodied in one computer system or separate computer systems. Accordingly, the first interface 120 and the second interface 122 may be internal interfaces of a computer or external interfaces that connect multiple computers. Communication between the first device 102 and the second device 104 may involve the first device 104 transmitting data to the second device 102 whereinafter the second device 102 processes, or otherwise modifies the data, and returns the processed data to the first device 104.

[0013] In some embodiments, the first device 102 may transmit data (e.g., a process request and associated data mentioned above) to the second device 104 via the first interface 120 which may comprise a full duplex connection (link). As the term is used herein, a full duplex connection refers to a connection in which data may be transmitted between the first device 102 and the network link adapter 106 in both directions simultaneously without affecting the speed of either data transfer. For example, Infiniband 1x, 2x, 4x, and 12x links comprise full duplex connections. Additional full duplex connections may comprise switched Ethernet over twisted pair wire or optical fiber, and Fibrechannel. Likewise, local computer interconnects such as HyperTransport and PCI-Express may comprise

full duplex connections. Via the first interface 120, the network link adapter 106 may receive data from the first device 102, whereinafter the network link adapter 106 may examine destination (target) information provided with the data and forward the data to the designated destination.

[0014] In some embodiments, the first device 102 may transmit data to any of a variety of devices (e.g., the second device 104) coupled to the network link adapter 106. Accordingly, devices coupled to the network link adapter 106 may each have an assigned address space (e.g., one or more addresses) defined by the network link adapter 106. Therefore, the destination information transmitted with data from the first device 102 to the network link adapter 106 may comprise an address assigned to the address space of the second device 104.

[0015] The network link adapter 106 (e.g., a Mellonox® MT23108 adapter) may be configurable to operate in multiple modes. As shown in Figure 1, the network link adapter 106 may comprise mode logic 114 that receives a mode signal. In some embodiments, the mode signal may be a user-controlled signal that permits the network link adapter to operate in one of multiple modes. Alternatively, the mode signal may be implicit in an address associated with the data received by the network link adapter 106. For example, in a first mode, the network link adapter 106 may forward the data to the second device 104 via the second interface 122 which may comprise a half duplex connection. As the term is used herein, a half duplex connection permits data to be transmitted in one direction at a time between the network link adapter 106 and the second device 104. For example, PCI, PCI-X, AGP and legacy busses such as ISA and EISA may comprise half duplex connections.

[0016] In a second mode, the network link adapter 106 may function to forward the data to a memory 108 (e.g., a DDR memory) via a memory interface 124 when the data targets a predetermined device. For example, in some embodiments, when data is intended for the second device 104 (as determined by destination information included with the data), the network link adapter 106 may be configured to forward the data to one or more predetermined addresses (real or dummy addresses) of the memory 108 via the memory interface 124. Therefore, when the first device 102 sends data that is intended for the second

device 104, the network link adapter 106 may redirect the data away from the second interface 122 to the memory interface 124. The second interface 122 (which may only transmit data in one direction at a time) is thereby available to transmit data from the second device 104 to the network link adapter 106 as will later be described.

[0017] As shown in Figure 1, the system 100 also may comprise a monitoring unit 110 coupled to the memory interface 124. The monitoring unit 110 is configurable to detect when data forwarded to the memory 108 was originally intended for the second device 104. For example, in some embodiments, the monitoring unit 110 may comprise an address detector 112 that detects when data is written to the one or more predetermined addresses described above. In response to detecting that data is being written to the one or more predetermined addresses, the monitoring unit 110 may intercept the data being written to the memory 108 and forward the data to the second device 104. Alternatively, the monitoring unit may read the data written to the one or more predetermined addresses and forward the data to the second device 104. If the address detector 112 does not detect that the data is associated with the one or more predetermined addresses, the data written to the memory 108 is not forwarded to the second device 104.

[0018] In response to receiving the data from the monitoring unit 110, the second device 104 may process the data and return the processed data to the first device 102 via the second interface 122. Therefore, the system 100 allows the second device 104 to receive data from the first device 104 via the alternate route (shown by the dashed line) while sending data to the first device 102 via the second interface 122. Without re-directing the data intended for the second device 104 as described above, the second interface 122 may act as a bottleneck that permits sequential rather than simultaneous communication between the first device 102 and the second device 104 (*i.e.*, the second device 104 would be prevented from outputting data to the first device 102 via the second interface 122 until all the data from the first device 102 has been received and the second interface 122 is available). Therefore, the speed of communication of system 100 may be increased (as much as doubled) by simultaneously providing data (*e.g.*, a

processing request) from the first device 102 to the second device 104 through the monitoring unit 110 while allowing the second device 104 to send data (e.g., processed data) back to the first device 102 data via the second interface 122, network link adapter 106, and the first interface 120.

[0019] In some embodiments, the system 100 may operate as, at least, part of a scalable graphic rendering system wherein the first device 102 is a network switch capable of dynamically communicating with a plurality of remote computer systems which work together in a coordinated manner to render computer generated imagery. In such embodiments, the second device 104 may process one or more sets of intermediate image data (*i.e.*, data that requires further processing/formatting to create an image) for a local computer that may be coupled to the second device 104 and/or one or more remote computers that may be coupled to the first device 102.

[0020] For example, a first set of intermediate image data may be transmitted from a remote computer to the first device 102 via a network connection. The first set of intermediate image data may then be transmitted from the first device 102 to the network link adapter 106 via the first interface 120. Thereafter, pre-established routing information, header information and/or addressing information associated with the incoming first set of intermediate image data may cause the network link adapter 106 to deliver the first set of intermediate image data to one or more pre-determined addresses in memory 108 via the memory interface 124.

[0021] As previously described, the monitoring unit 110 may observe the data on the memory interface 124 and forward the data to second device 104 via an alternate interface (the alternate route) that does not interfere with the operation of second interface 122. The second device 104 may combine the first set of intermediate image data with a second set of intermediate image data received from a local computer coupled to the second device 104. The combination of the first and second sets of intermediate image data may produce new image data that may be displayed by the local computer or one of the remote computers. Alternatively, the new image data may comprise intermediate image data that is subject to further processing/combining in the local computer or one of the remote computer systems prior to final display. In at least some embodiments, the

remote computer system that provided the first set of intermediate image data to second device 104 is not the same as the remote computer system that may receive the new image data from the second device 104.

[0022] Figure 2 illustrates a flow of data in the system 100 of Figure 1 in accordance with embodiments of the invention. As shown in Figure 2, the flow of data (shown in bold) may be divided into "PHASE 1," "PHASE 2" and "PHASE 3." During PHASE 1, data is transmitted from the first device 102 to the network link adapter 106 via the first interface 120. If the second device 104 is the desired recipient of the data and the network link adapter 106 is configured to operate in the second mode described above, the data is forwarded to one or more predetermined addresses of the memory 108. During PHASE 2, the monitoring unit 110 detects that data is being written to the one or more predetermined addresses and forwards the data to the second device 104. The second device 104 may then process the data as requested by the first device 102. During PHASE 3, the processed data is returned to the first device 102 via the second interface 122, the network link adapter 106, and the first interface 120. In at least some embodiments, the PHASES described above may occur substantially simultaneously such that the first device 102 may still be outputting a data process request to the second device 104 while receiving processed data back from the second device 102.

[0023] Figure 3 illustrates a method 200 in accordance with embodiments of the invention. As shown in Figure 3, the method 200 may comprise configuring a network link adapter to write to a predetermined memory address when data received from a first device is intended for a second device (block 202). At block 204, the data traffic to the memory is observed. If the network link adapter is writing data to the predetermined memory address, as determined at block 206, the data may be forwarded to the second device (block 208). Otherwise, the method 200 waits for the adapter to write data to the predetermined address (block 206).

[0024] The above discussion is meant to be illustrative of the principles and various embodiments of the present invention. Numerous variations and modifications will become apparent to those skilled in the art once the above

disclosure is fully appreciated. It is intended that the following claims be interpreted to embrace all such variations and modifications.